

SIGNAL PROCESSING for ION MOBILITY SPECTROMETERS

by S. Taylor, M. Hinton, R. Turner, Graseby Dynamics Ltd, Odhams Trading Estate, St. Albans Road, Watford, Herts, WD2 5JX, UK

N95- 33289

1995726868

ABSTRACT

Signal processing techniques for systems based upon Ion Mobility Spectrometry will be discussed in the light of 10 years of experience in the design of real-time IMS. Amongst the topics to be covered are compensation techniques for variations in the number density of the gas - the use of an internal standard (a reference peak) or pressure and temperature sensors. Sources of noise and methods for noise reduction will be discussed together with resolution limitations and the ability of deconvolution techniques to improve resolving power. The use of neural networks (either by themselves or as a component part of a processing system) will be reviewed.

1. BACKGROUND

Graseby Dynamics has been developing ion mobility spectrometers for over 15 years. We have produced systems that are capable of automatic detection of a number of chemical species :

Chemical warfare agents - for example nerve and blister agents

Drugs - heroin, cocaine, amphetamines etc.

Explosives - nitro-glycerine, TNT, RDX etc.

Hazardous chemicals - TDI , dimethyl sulphate etc.

2. OBJECTIVE

This paper will not consider how to present the sample to the analyser but only how to deal with the signals generated by an IMS after the sample has been induced.

The objective of the signal processing is to identify and quantify the analyte.

Most of the work that Graseby has done has been aimed at automatic identification - the spectrometers have to analyse, identify and quantify either completely unattended or at least without operator intervention.

Figure 1 depicts the overall signal processing system.

3. FUNDAMENTALS

What sort of signals ?

Ion currents - generally below the nanoamp region (though some ion sources may exceed this).

Time -varying and repetitive with repetition times between a few milliseconds to a few hundred milliseconds. (This assumes a single ion gate system with direct time of travel analysis).

What sort of noise ?

3.1 Statistical noise as a consequence of the ion numbers - generally not a significant factor:

Corresponds to a standard deviation of about 0.001 % at the 1pA level.

(save for the rather special case of low activity alpha sources which can be noisy as a result of the formation of bursts of ions along each alpha track.)

3.2. Classical noise sources in the amplifying system

Johnson (thermal) noise tends to predominate over other classical noise sources.

It is advantageous to employ a current-to-voltage converter as the first stage. This minimises effects of capacitance on the bandwidths since the input is a virtual earth.

The amplifying system needs a reasonable bandwidth to accurately amplify time varying signals.

Figure 2 depicts the measured bandwidth of a typical IMS spectrum - compared to the anticipated bandwidth of Gaussian shaped peak of the same half-height width.

Biggest value in the feedback resistor (R) of the current-to-voltage converter gives the best S/N since noise increases as square root of resistance whilst the gain is directly proportional to R.

3.3 Electrical noise

There is a need for efficient screening of the signal path both inside and outside the spectrometer itself - in CAM a significant amount of electrical noise can be generated from the various signals flying around the system - one of the LCD driver signals (10 volts peak-to-peak) was found to feed through a small gap in the screening and cause a significant degradation in S/N.

Poor screening within the drift tube can lead to noise on the field defining electrodes being passed onto the collector electrode.

In complete instrument systems there can be a galaxy of noise generators - digital lines, switch mode power supplies etc. and care must be taken to avoid other noise routes into the signal path. For example noise on the reference used for analogue to digital conversion can reflect directly onto the signal.

3.4 Mechanical noise

There are two main areas where mechanically induced noise may arise :

- a) The ion collector / collector screen assembly within the drift tube.

and b) The electrical leads taking the signal to the current-to-voltage converter.

The collector screen serves to shield the collector from variations in electric field - including that caused by an approaching cloud of ions. It is held in close proximity to the collector electrode but there is a fixed potential between the two. This arrangement is very like that of a condenser microphone. Movement of one component with respect to the other induces currents into the ion collector and unless great care is taken in order to achieve mechanical rigidity the arrangement is very susceptible to vibration.

If, for engineering reasons, the current-to-voltage converter cannot be mounted immediately beside the collector then a length of co-axial cable may have to be used. Mechanical vibration of the cable can induce charge separation between the inner and outer conductors of the cable by triboelectric and piezoelectric effects. This induced current is amplified at the current-to-voltage converter and significant microphony can result. Careful selection of the type of co-axial cable is required along with design steps to reduce the level of the vibration.

4. NOISE REDUCTION

4.1 Filtering

Clearly one needs to filter out as much noise as possible without distorting the signal significantly. Analogue filters with roll-off characteristics completely matching the frequency content of the IMS signal would be the ideal to apply. In practice, filters of standard types (Butterworth, Bessel, Chebyshev etc) tend to be applied. They provide the engineering compromise between the ideal and that achievable without adding spurious effects e.g. ringing and overshoot. In any event the frequency characteristic of an IMS product ion peak varies with the drift time associated with it due to peak broadening effects. The ideal filter would have to have a constantly varying frequency response.

We have applied both analogue and digital filtering methods. One should always apply analogue filter methods fully in addition to any further digital processing. There is a very large range of filter techniques available. We tend to favour using a Bessel filter of 5th order. The pass/stop transition region we select depending on the application. For instance, for our high temperature IMS systems we need a higher bandwidth to cater for the sharper peaks caused by the shorter travel time at the elevated temperatures.

Filtering methods introduce phase delays in the filtered signal. These can be significant and should be considered if the ultimate in reduced mobility calculation is required. One of the reasons for selecting a filter with a Bessel response rather than a Butterworth response is that the former gives a constant phase delay for any frequency within its pass band. Typically, the calculated phase shifts will be of the order of 100-200 μ s and should be taken into account when calculating reduced mobility figures from first principles.

4.2 Averaging

The time-coherent addition of repetitive ion mobility spectra is very effective at improving the signal/noise ratio. In theory, for an average of N spectra, a \sqrt{N} improvement in signal/noise ratio should result. The technique may be considered as a time-domain filter with a very narrow pass band i.e. it will only pass a signal that is in strict time coherence. In practice, the improvement is usually slightly less than \sqrt{N} . Some noise sources may have a small time

coherence for a small N e.g. a pump vibration induced noise may have a period similar to the repetition rate of the IMS. In this case a random time delay can be inserted between each trigger of the IMS to interrupt the time coherence between the IMS signal and the rotation rate of the pump motor.

The disadvantage of averaging is that fast changes in the spectrum will be smoothed out. This may lead to an apparent reduction in GC resolution in GC-IMS. The smoothing action can be vitally important if the effect is to delay warning a soldier that he should put on his gas mask. Dynamic averaging can be applied where the spectra are checked after an average of 4, say, if there is no significant change i.e. none above the noise, then the averaging process can add in the previous 4 spectra and recheck and so on. Such techniques attempt to optimise the speed of response to a signal/noise ratio that gives a trustworthy measurement.

A similar, but more comprehensive, approach could be adopted in GC-IMS if all the spectra were stored separately for later processing. This might entail conducting intelligent averaging whereby, if the spectra had fast-changing aspects, say, as a large sample pulse eluted from the column, then fewer averages or a rolling average could be conducted for this part of the time record. If the IMS peaks were small then more averaging would be conducted to increase the signal/noise ratio. In practice such advanced techniques have not proved necessary since with taking, say, averages of 16 spectra it is found that 10 or more data points may be acquired per GC peak.

5. RESOLUTION

The resolution of an IMS signal places fundamental limits on the ability of the IMS system to identify individual ion species. The resolution may be measured in a number of ways, but is essentially the ability of the instrument to separate and quantify packets of ions of slightly different mobility. In signal processing terms it relates to the width of the IMS peaks and the time separation between them.

IMS peak broadening is caused by a number of effects. Most notable amongst these are diffusion Gaussian broadening and space charge effects. For very small ion peaks (i.e. a small number of ions) the space charge effects may be ignored and the diffusion broadening then places an upper limit on the IMS resolution. In the absence of space charge effects, integration of the Boltzmann equation results in the resolution being proportional to the square root of the drift voltage. In the limit of increasing drift voltage the half-height width of the peak will tend towards the width of the gate pulse and then the resolution becomes proportional to the inverse of the drift voltage.

Figure 3 shows the variation in peak width with ion density (peak height) for the experiment where the ^{63}Ni ion source was placed at differing distances from the ion gate. This gives an indication of space charge effects on the peak widths.

6. RESOLUTION ENHANCEMENT

Deconvolution is a signal processing technique used to enhance the resolution by removing the peak broadening process suffered by the ion packet as it travels down the drift tube. In its simplest form the broadening process may be viewed as a transfer function acting on the signal (gate open pulse) applied to the system.

$$y = D * x$$

where y is the output signal,
x is the input signal
& D is the peak broadening operator

If we know this function perfectly (i.e. its impulse response is known and that there is no noise in the output) then the deconvolution process is to apply the inverse of the transfer function to the signal:-

$$x = D^{-1} * y$$

This may be achieved by the use of Fourier transforms.

However, when noise, n, is present, that is;

$$y = Dx + n$$

and we must determine:-

$$x = D^{-1}(y - n)$$

$D^{-1}n$ can be large for parts of the frequency spectrum and this simple method then amplifies noise in the signal and is unusable. A further problem is that small errors in determining D can lead to large errors in D^{-1} leading to spurious results and noise amplification.

There are many deconvolution schemes published in the literature. At Graseby Dynamics we have had some success with two deconvolution schemes. Both require a reasonable estimate of the impulse response of the system transfer function, but they have been shown to be substantially less prone to producing artefacts due to noise in the original signal.

A) Constrained Iterative Deconvolution

This uses an iterative approach, as described by Schafer et al¹, with the constraint that no part of the resultant output (deconvolved spectrum) of each iteration can be less than zero. This corresponds with the physical situation that charge of incorrect polarity cannot arrive at the collector electrode.

Figures 4 and 5 show a typical iteratively deconvolved spectrum

The resolution may be enhanced by a factor of about 3. The peak magnitude information can be unreliable after applying deconvolution. Peak height/area measurements should then be conducted on the original spectrum and the deconvolved spectrum be used solely for identification.

B) Maximum Entropy Deconvolution

Graseby Dynamics has used the services of Maximum Entropy Data Consultants Ltd² (MEDC) to conduct their deconvolution method on some trial spectra. Again an estimate has to be made of the impulse response of the system transfer function. For this and other studies we have found that a simple Gaussian shape is a reasonable first approximation, but a better approach is to take the spectrum produced by a single ion species e.g. for positive ion mode

that of acetone in an otherwise clean system. Since the stimulus of the system is almost an impulse, then the output (i.e. the acetone RIP) is the response to this and can be taken as the transfer function. MEDC applied their method, using an acetone RIP as the transfer function and obtained very good results. The resolution may be enhanced by a factor of about 8.

Figures 6 and 7 show the same spectrum as that in figure 4, but now deconvolved by the maximum entropy method.

As can be seen the MEM deconvolved peaks have the best resolution and the noise level is also reasonable. MEDC claim that the principles they incorporate in their processing are the only logical ones which should be applied to deconvolution. Further, they claim their processing allows for a rigorous, statistically based estimation of ion peak magnitudes in the deconvolved spectrum. That is, they can measure a peak magnitude and provide a confidence figure for the value's accuracy.

Of course, there has to be a disadvantage to such a powerful technique - it is very computationally intensive. It requires several hundred to a few thousand FFTs to perform a single deconvolution. It may not yet be a real time technique, but it may find use in those laboratory applications where resolution is paramount.

7. IDENTIFICATION

Having obtained a spectrum of low noise and high resolution, the next problem is to identify ion species by means of accurate reduced mobility measurements. Two methods may be adopted:-

- a) Use of a standard reference compound. In which the ratio of the target compound ion drift time to that of a known reference ion gives the reduced mobility.
- b) Use of pressure, temperature and drift voltage measurements. In which, these parameters, the drift length and the target compound ion drift time are used to calculate the reduced mobility.

Both methods have merits and disadvantages. In the reference compound method, the selection of the compound and the manner of its introduction to the system are vital. Many reduced mobilities are concentration and temperature dependent.

To measure the reduced mobility accurately from first principles requires an accurate knowledge of the ion drift time, the drift length, the drift voltage and the IMS internal temperature and pressure. Further, the drift field must be perfectly regular for the entire drift length and the influence of end effects and signal processing delays must be known. In practice the instrument will be calibrated using a compound producing a known, stable ion species, with an internationally accepted reduced mobility figure. The steps being taken to choose standard compounds are to be commended.

The calibration provides an instrument constant which includes the unknowns listed above. Thereafter, the drift voltage, drift tube temperature and pressure need to be accurately measured. An example of a most-likely error bound calculation is given below. The error figures quoted for each measurement are typical of what may be achieved in a commercial instrument.

$$Ko = \xi \frac{1}{t} \frac{P}{T} \frac{1}{V}$$

where ξ = a instrument constant determined using a reference compound

t = drift time

P = pressure

T = temperature

V = drift voltage

$$\left(\frac{\delta Ko}{Ko}\right)^2 \approx \left(\frac{\delta t}{t}\right)^2 + \left(\frac{\delta P}{P}\right)^2 + \left(\frac{\delta T}{T}\right)^2 + \left(\frac{\delta V}{V}\right)^2$$

With careful design:-

$$\left(\frac{\delta Ko}{Ko}\right)^2 \approx \left(\frac{10}{10000}\right)^2 + \left(\frac{5}{1000}\right)^2 + \left(\frac{0.15}{300}\right)^2 + \left(\frac{1}{1000}\right)^2$$

$$\text{Giving } \left(\frac{\delta Ko}{Ko}\right) \approx 0.5\%$$

For most applications the selection of a small, low power pressure transducer means that the errors in this measurement tend to dominate the overall error. For this not to be the case one would need to measure the absolute pressure to +/- 1mbar.

Having decided by either method where to look for ion peaks in the spectrum, the next process is to determine the presence and position of peaks. This may be achieved by means of peak searching routines followed by the application of a set of decision rules, i.e. if a peak appears within a time/mobility window then it must be compound, X. Our experience is that the set of decision rules often becomes increasingly complicated throughout the life of the instrument as new fields-of-use and interferences are encountered. We are now using combined positive and negative ion mode spectral data to identify compounds, collected simultaneously using twin drift tube systems, to enhance interferent rejection.

Increasingly, neural networks are being applied to conduct the identification function. The result of an experimental neural network is given in tables 1 & 2 below. The network was fed the spectra of 4 different compounds over a range of concentrations and the spectra of the clean system (no sample vapour present) Table 1 shows the values in the 5 output elements once the network has been trained on this set of data. Not surprisingly the output element associated with each compound gives a high value when offered the spectrum of that compound. More interestingly the network also performed reasonably well with the separate test set of spectra of the same compounds as shown in table 2.

Table 1 Neural Network Results (Training set)

1 (A)	Output element number (Compound ident)				Compound
	2 (B)	3 (C)	4 (D)	5 (RIP)	
1	0	0	0	0	A
1	0	0	0	0	A
1	0	0	0	0	A
0	1	0	0	0	B
0	1	0	0	0	B
0	1	0	0	0	B
0	0	1	0	0	C
0	0	1	0	0	C
0	0	1	0	0	C
0	0	0	1	0	D
0	0	0	1	0	D
0	0	0	1	0	D
0	0	0	0	1	RIP
0	0	0	0	1	RIP
0	0	0	0	1	RIP

RIP refers to the spectrum from a clean system (no sample vapour).

Table 2 Neural Network Results (Test set)

1 (A)	Output element number (Compound ident)				Compound
	2 (B)	3 (C)	4 (D)	5 (RIP)	
1	0	0	0	0	A
0.9	0	0	0	0.1	A
1	0	0	0	0	A
0	1	0	0	0	B
0	0.8	0	0	0.1	B
0	1	0	0	0	B
0	0	1	0	0	C
0	0	0.8	0	0.26	C *
0	0	1	0	0	C
0	0	0	1	0	D
0	0	0	1	0	D
0	0	0.3	0.9	0	D
0	0	0	0	0.9	RIP
0	0	0	0	1	RIP
0	0.1	0	0	0.9	RIP

Note that the result for the test labelled marked (*) was for a low concentration of compound, C, such that the RIP was near its full height. It is not surprising that the neural network considered that this spectrum was similar to that for a clean system. With a larger training set this, and other deviations from the expected result may have been avoided.

The above results were obtained using digitised spectra which were scaled for pressure and temperature (drift voltage assumed constant). The complete spectrum was offered to the

network in each case - not just peak positions and heights. This method can distinguish compounds that give very similar spectra. It is also found to give better interferent rejection and can be trained to reject specific interferent compounds.

8. QUANTIFICATION

A typical IMS spectrum consisting of 16 averages has a figure for the ratio (clean-system RIP/noise level) of approximately 100. This implies a dynamic range of about 2 orders of magnitude which could be extended by reducing the noise. A quiet system and with the spectra acquired using hundreds of averages the dynamic range may be extended to 3 orders of magnitude. Of course, only a small section of this range, usually at low concentrations, may be considered linear. At high concentrations functional forms for the quantitation need to be applied to deal with the appearance of dimers and trimers and they may also aid in overcoming some of the saturation effects. A functional form may be represented:-

$$\text{Measured value} = f(M, D, RIP)$$

where M = monomer peak height

D = dimer peak height

RIP = RIP peak height

These may take many forms, depending on the instrument and ion chemistry. A typical form might be:-

$$\text{Measured value} = M + D \text{ for } RIP > \text{Trip value}$$

$$\text{Measured value} = M + D + \alpha \frac{D}{M} \text{ for } RIP < \text{Trip value}$$

where α is a constant derived from experiment

In calibrating an IMS system at Graseby Dynamics we generally use look up tables to convert from the measured value to a concentration value. This allows for greater flexibility than trying to use a single equation to fit the response curve over the entire dynamic range of the instrument.

Peak areas may be used instead of peak heights and may be preferred if the increase in noise from the summing of several digitised samples within the spectrum is less than the increase in the signal. This suggests that one should only measure the central area of the peak, out to about the half-height position.

REFERENCES

1. R. W. Schafer, R. M. Mersereau & M. A. Richards
Proc. IEEE vol 69 No4 p432 (April 1981)
2. Maximum Entropy Data Consultants Ltd, St. John's College, Cambridge, CB2 1TP, U.K.

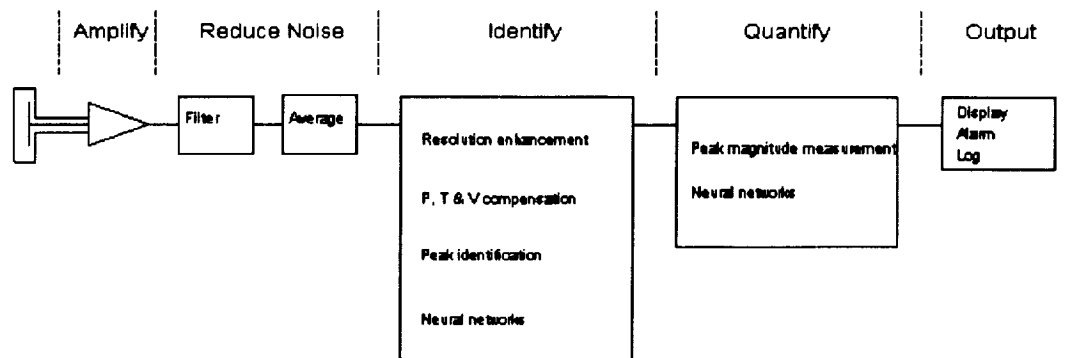


Figure 1. Overall signal processing path

ORIGINAL PAGE IS
OF POOR QUALITY

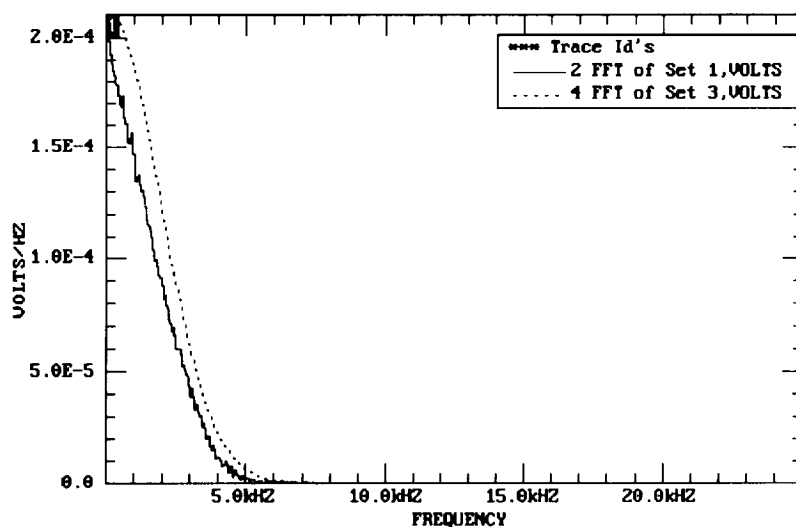


Figure 2. Typical frequency spectrum of an IMS signal (solid) and that of a calculated Gaussian peak of similar half-height width (dashed)

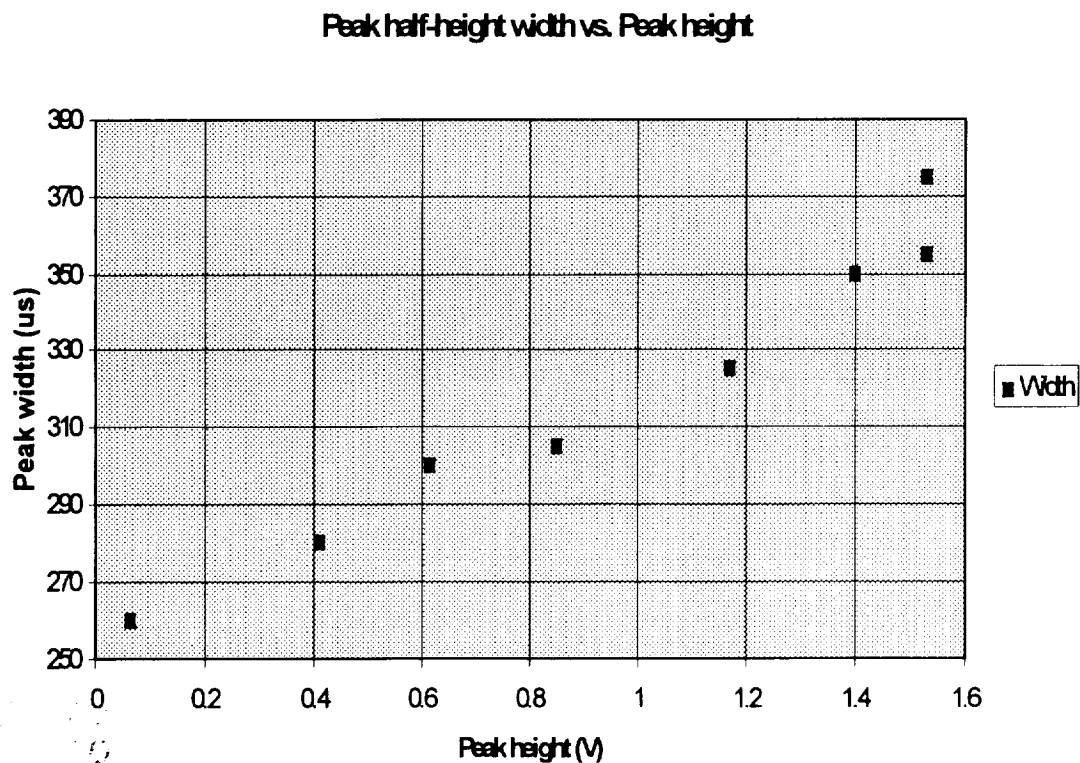


Figure 3. Ion peak half-height width versus ion peak height when the source-to-gate distance is varied.

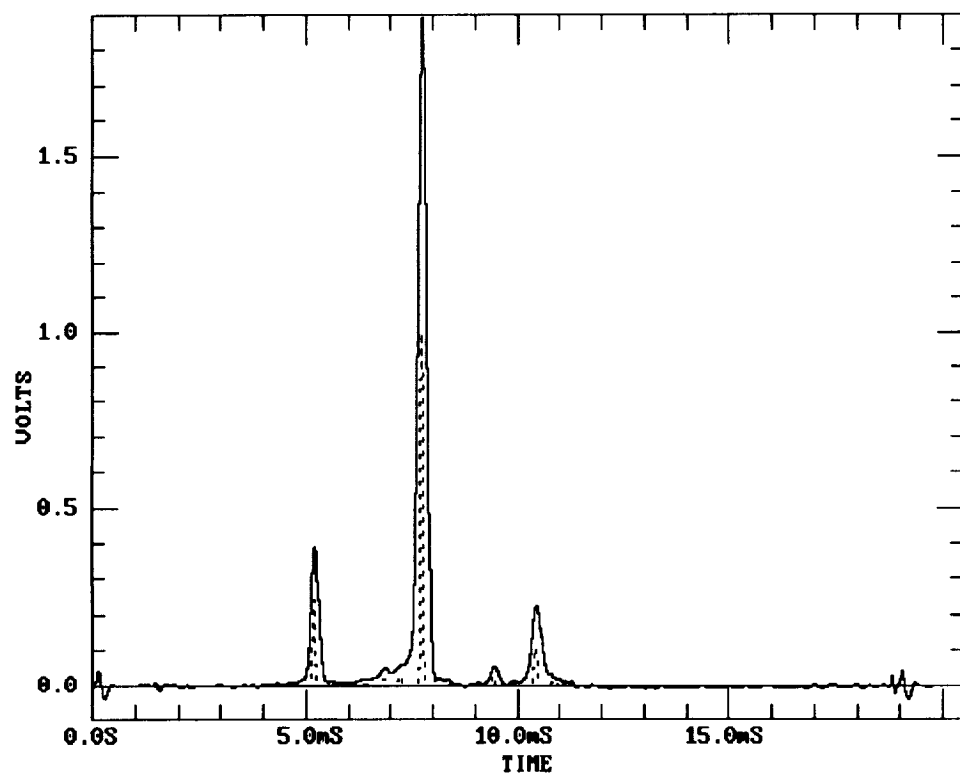


Figure 4. Typical original (solid) and iteratively deconvolved (dashed) spectrum

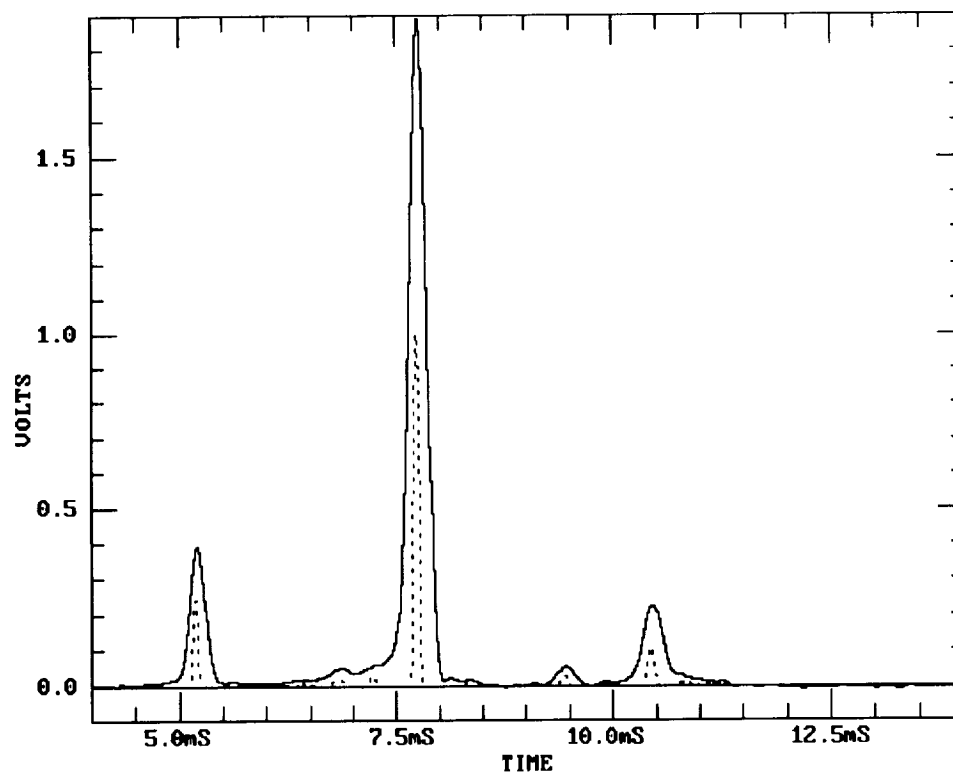


Figure 5. Spectrum from figure 4 on expanded timescale

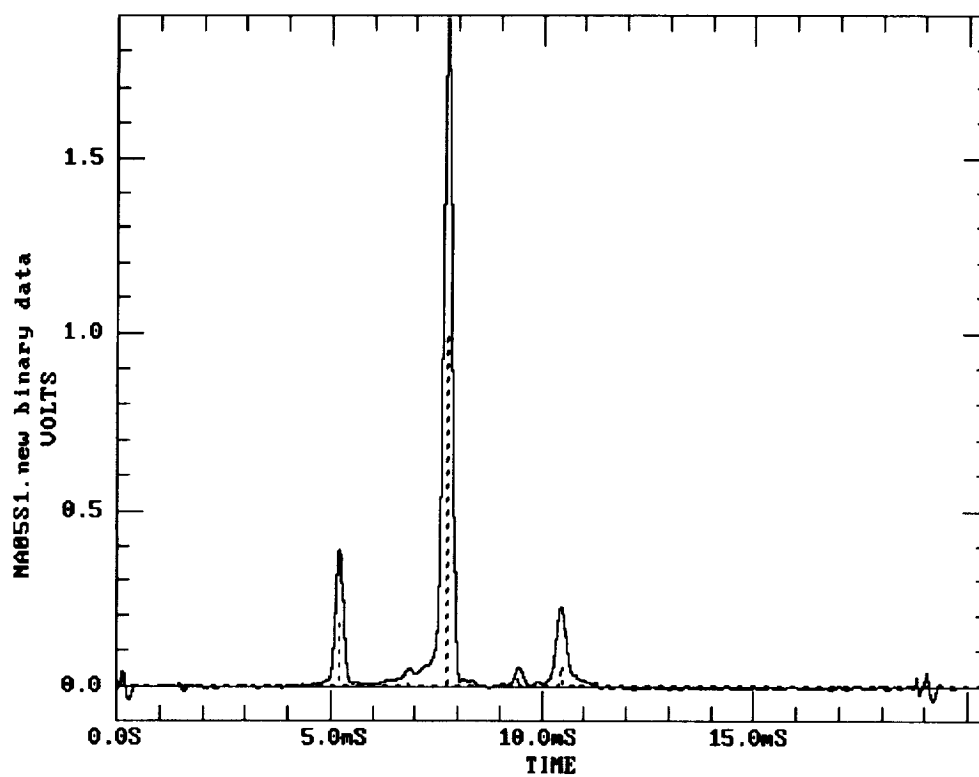


Figure 6. Typical original (solid) and maximum entropy deconvolved (dashed) spectrum

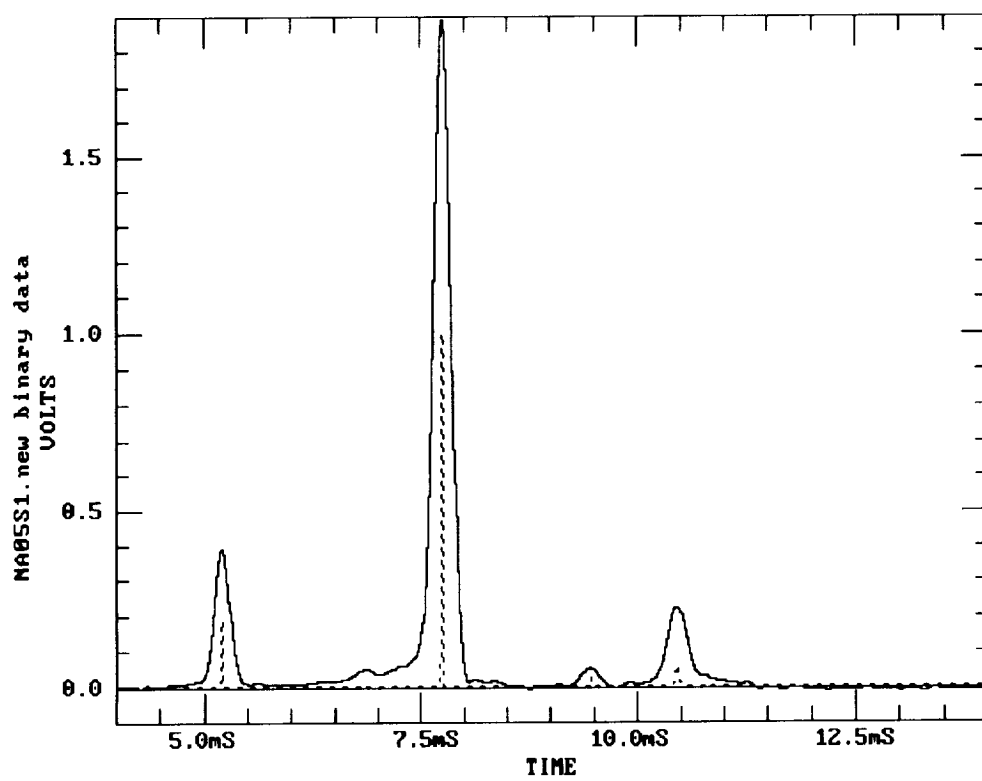


Figure 7. Spectrum from figure 6 on expanded timescale